PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Heterogeneity in District-Level Transmission of Ebola Virus Disease during the 2013-2015 Epidemic in West Africa

Fabienne Krauer[1]*, Sandro Gsteiger[1], Nicola Low[1], Christian H. Hansen[2], Christian L. Althaus[1]

1 Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland, **2** MRC Tropical Epidemiology Group, London School of Hygiene & Tropical Medicine, London, United Kingdom; Mwanza Intervention Trials Unit, National Institute for Medical Research, Mwanza, Tanzania

* fabienne.krauer@ispm.unibe.ch

Citation: Krauer F, Gsteiger S, Low N, Hansen CH, Althaus CL (2016) Heterogeneity in District-Level Transmission of Ebola Virus Disease during the 2013-2015 Epidemic in West Africa. PLoS Negl Trop Dis 10(7): e0004867. doi:10.1371/journal.pntd.0004867

Editor: Andrew Fenton, University of Liverpool, UNITED KINGDOM

Received: January 7, 2016

Accepted: June 30, 2016

Published: July 19, 2016

Data Availability Statement: EVD case data are available from the Ebola data and statistics database of WHO (http://apps.who.int/gho/data/node.ebola-sitrep.ebola-country?lang=en). Subnational population density estimates are available from the censuses (http://www.stat-guinee.org/index.php/result-prelim-rgph3, http://www.lisgis.net/page_info.php?7d5f44532cbfc489b8db9e12e44eb820=MzQy, http://www.statistics.sl/2004_population_and_housing_census.htm). Subnational estimates of socio-demographic factors are derived from DHS

Abstract

The Ebola virus disease (EVD) epidemic in West Africa in 2013–2015 spread heterogeneously across the three hardest-hit countries Guinea, Liberia and Sierra Leone and the estimation of national transmission of EVD provides little information about local dynamics. To investigate district-level transmissibility of EVD, we applied a statistical modelling approach to estimate the basic reproduction number (R_0) for each affected district and each country using weekly incident case numbers. We estimated growth rates during the early exponential phase of the outbreak using exponential regression of the case counts on the first eight weeks since onset. To take into account the heterogeneity between and within countries, we fitted a mixed effects model and calculated R_0 based on the predicted individual growth rates and the reported serial interval distribution. At district level, R_0 ranged from 0.36 (Dubréka) to 1.72 (Beyla) in Guinea, from 0.53 (Maryland) to 3.37 (Margibi) in Liberia and from 1.14 (Koinadugu) to 2.73 (Western Rural) in Sierra Leone. At national level, we estimated an R_0 of 0.97 (95% CI 0.77–1.18) for Guinea, 1.26 (95% CI 0.98–1.55) for Liberia and 1.66 (95% CI 1.32–2.00) for Sierra Leone. Socio-demographic variables related to urbanisation such as high population density and high wealth index were found positively associated with R_0 suggesting that the consequences of fast urban growth in West Africa may have contributed to the increased spread of EVD.

Author Summary

Since 2013, a major epidemic of Ebola virus disease (EVD) spread throughout three countries in West Africa for about two years. Its scale was unprecedented since the Ebola virus was first identified. Almost all districts of Guinea, Liberia and Sierra Leone were affected, but some areas observed much larger outbreaks than others. In this study, we offer insights into the geographical variation of EVD transmissibility. We estimated the epidemic growth rate for each subnational unit and used the results to calculate the basic

datasets available at the DHS program website (http://www.dhsprogram.com/).

Funding: CLA received funding through an Ambizione grant from the Swiss National Science Foundation (project 136737). SG was supported by the Epidemiology and Mathematical Modelling for Infectious disease Control (EpideMMIC) project (Swiss National Science Foundation project number 320030_135654). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

reproduction number R_0, which indicates how many secondary infections arise on average from one primary case at the beginning of an outbreak. We found that the transmissibility of EVD differed within the three countries and we identified areas of high initial transmission. We also show that socio-demographic factors related to urbanisation and crowding might have been acting as drivers of EVD spread. These findings confirm that early containment of outbreaks in large communities is crucial for the prevention of epidemics as seen in West Africa during 2013–2015.

Introduction

The Ebola virus disease (EVD) outbreak that started in December 2013 in Guinea developed into the largest EVD epidemic ever observed. There has been some discussion about the geographical heterogeneity of disease transmission in the three hardest hit countries in West Africa [1–8], but other studies have not considered this effect in their analysis. An epidemic of this scale has an intrinsic multi-level structure and national epidemic curves are always an overlay of local outbreaks [9–11]. Consequently, the estimation of national transmission offers little information about subnational dynamics. The availability of district-level data provides a unique opportunity to investigate local disease transmission during the 2013–2015 epidemic. We may also use these data to quantify population-level risk factors for EVD transmission. Demographic or behavioural factors such as crowding and high population density, low socio-economic status (SES), unsafe burials or poor sanitation but also climate effects might have contributed to enhanced EVD transmission in West Africa, but these effects have been assessed only in a limited number of studies [1,7,11,12]. In the context of recurrent infectious diseases, increased human mobility and globalisation the quantification of subnational spread of EVD and the investigation of factors related to the spread might provide insights that could improve epidemic management in the future [5].

The most frequently used parameter for quantifying transmissibility is the basic reproduction number R_0, which describes the average number of secondary infections generated by a primary case during the initial phase of an outbreak when the population is completely susceptible and no control measures have been employed [13]. A common approach to the estimation of R_0 consists of fitting mathematical transmission models to observed outbreak data. In the West African EVD epidemic, several mathematical modelling studies have described the variation in R_0 between [6,10,14–19] and within [20–25] the three countries. Most studies found estimates of R_0 ranging between 1.5 and 2.5, agreeing with results from models of earlier outbreaks ranging from 1.4–4.7 (summarised in [26]). A drawback of mechanistic models is the requirement of a large number of parameters, which is problematic with sparse data [27]. As an alternative to mechanistic models, R_0 can be inferred from the generation time and the intrinsic growth rate r (sometimes also referred to as Λ) during the early, exponential epidemic phase [13]. The generation time is the average interval between infection of a primary and a secondary case. The growth rate r is defined as the per capita change in case numbers per time unit. It can be calculated directly from the observed empirical data by estimating the slope of the natural log-transformed cumulative case counts over time with a linear regression model [28,29] or by fitting an exponential or sigmoid curve to incident case counts [27,30]. The use of incident case data is generally preferred over cumulative case data because the individual observations are statistically independent [1]. Ignoring the dependence of measurement error in cumulative cases can lead to over-optimistic standard errors and thus underestimation of uncertainty in epidemic growth [31]. This statistical approach is particularly useful for the

assessment of smaller outbreaks with sparse data [27], as observed in certain districts during the West African EVD epidemic. The analysis of local transmissibility should take into account the hierarchical nature of subnational data, because these cases are not completely independent but arise from a common population i.e. country. Such an epidemic structure can be approached with a mixed effects model, which explicitly allows for clustering.

In this study, we used district-level case data from the Ebola epidemic in West Africa and implemented a statistical modelling approach to calculate the epidemic growth at country and at district level. We then estimated the transmissibility of EVD expressed as R_0 for Guinea, Liberia and Sierra Leone and their 53 affected districts. Finally, we explored the relationship between R_0 and socio-demographic variables as potential drivers of EVD transmission at a population-level.

Methods

EVD case data

We used data on weekly incident cases of EVD in each subnational unit (here called 'district') of Guinea (préfécture), Liberia (county) and Sierra Leone (district), which are available at the Ebola data and statistics website of the World Health Organization (WHO) [32]. These weekly case counts were aggregated from the patient database and are considered to be more reliable for the early phase of the epidemic than the situation reports issued by the Ministries of Health of the affected countries [33]. We calculated the total incident cases for each week as the sum of confirmed and probable incident cases. To take account of recurrent outbreaks we defined epidemic waves within each district. A wave was considered terminated after a period of 42 days with no new cases, corresponding to twice the maximum incubation period [34]. New cases arising after this period were considered as a new wave. To ensure a minimum number of data points in each district for the fitting process, we restricted our analysis to the first wave with three or more non-zero data points. Districts with fewer than three non-zero data points per wave were excluded. We restricted the time variable to the first eight weeks since onset of a wave in each district, corresponding to approximately four serial intervals [34]. This time period was considered long enough to have a sufficient number of time points and reduce initial stochasticity, but short enough to capture only the initial exponential phase and exclude the effects of control measures or natural attenuation of the epidemic. The comparison of the dates of epidemic onset and the opening dates of Ebola Treatment Units (ETUs, see [35,36]) showed that only 9% of all districts had a functioning ETU in the first eight weeks since onset of the selected wave.

Mixed effects model for the estimation of the epidemic growth rates

We assumed that the epidemic growth in individual districts varied both between countries and between districts within a country. We used a generalised linear mixed effects model (GLMM) comprising both fixed and random effects, which explicitly allows for clustering in the data [37]. Such a hierarchical model allows the mean values to vary between the different countries, but borrows information across the districts within a country. The weekly number of new infections c_{ijk} in district i in country j at time-point t_{ijk} was assumed to follow a Poisson distribution with a mean λ_{ijk} and was modelled with the logarithm as the link function:

$$c_{ijk} \sim Pois(\lambda_{ijk})$$

$$\ln(\lambda_{ijk}) = (\beta_{0j} + b_{0ij}) + (\beta_{1j} + b_{1ij}) * t_{ijk}$$

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_0^2 & \rho\tau_0\tau_1 \\ \rho\tau_0\tau_1 & \tau_0^2 \end{pmatrix}\right) \quad (1)$$

This means we assume country level intercepts β_{0j} and slopes β_{1j}, which are modelled as fixed effects. The district specific intercepts $\beta_{0j} + b_{0ij}$ and slopes $\beta_{1j} + b_{1ij}$ are assumed to be normally distributed around these country level average values. For simplicity, we assume the heterogeneity between the district-level intercepts (${\tau_0}^2$) and slopes (${\tau_1}^2$) as well as their correlation (ρ) to be the same for all countries. The model was fitted using maximum likelihood estimation. We then used the posterior means of the random effects (empirical Bayes means) to calculate the growth rates r for each district i in country j, given by

$$r_{ij} = \beta_{1j} + b_{1ij} \quad (2)$$

Basic reproduction number R_0

The expression relating r to R_0 is the inverse of the moment generating function of the generation time distribution, which uniquely identifies R_0 for a given r and generation time distribution [13]. If the generation time is known and if it follows a gamma distribution, R_0 can be calculated as

$$R_0 = \left(1 + \frac{r}{\beta}\right)^{\alpha} \quad (3)$$

where r is the intrinsic growth rate, α is the shape parameter and β is the rate parameter of the gamma distribution, respectively [13]. For known α and β, we can derive the uncertainty in R_0 from the uncertainty in the growth rate estimates r with the delta method [37]. The variance (Var) of R_0 is approximated as

$$\mathrm{Var}(R_0(r)) = \mathrm{Var}(r) * \left(\frac{\partial R_0(r)}{\partial r}\right)^2 \quad (4)$$

and the standard error (SE) is the square root of the variance. The lower and upper bounds of an approximate 95% confidence interval (CI) are then calculated as the estimate±1.96*SE derived with the delta method. The shape (2.59) and rate (0.17 per day) parameters of the serial interval were obtained by fitting a gamma distribution to the serial interval distribution reported by the WHO Ebola response team [34]. For simplicity, we did not consider the uncertainty in the estimation of α and β (which could be done by applying multivariate versions of the delta method).

The estimated epidemic growth rate might be influenced by the length of the time window under investigation. To examine the impact of assumptions about the exponential phase on our estimates, we performed a sensitivity analysis for the mixed effects model with time windows one to three weeks shorter (five to seven weeks) and longer (nine to eleven weeks) than the proposed eight weeks.

Associations between R_0 and socio-demographic factors

We used an ecological study design to quantify the potential impact of risk factors for EVD transmission at district level. We focused on variables that are strongly suspected to be drivers of infectious disease spread: population density, household density, low SES and poor sanitation. Population density was calculated using population sizes and surface areas derived from national censuses [38–40], the other variables were derived as district-level summary measures

from Demographic and Health Survey (DHS) datasets [41–43]. Household density was calculated as the number of persons per sleeping room. SES was approximated using the DHS wealth index, which is a composite score of household assets on a continuous scale transformed to a standard normal distribution for each country. An increasing score indicates a higher SES. For simplicity, we did not calculate a comparative wealth index score [44], which generally limits the comparability of the SES across countries, but we consider the use of the original score adequate for the purpose of this analysis. We used the average time to walk to the nearest water source as a proxy for the level of sanitation. The calculation of the summary estimates is described in detail in the supplement (S1 Text).

Statistical analysis

All analyses were carried out using Stata (StataCorp. 2013. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP). The mixed effects model was implemented using the meglm routine with the unstructured covariance option. The association between socio-demographic factors and transmissibility as measured by R_0 was quantified in a univariable linear regression model. All summary data are expressed as median and range or mean and standard deviation (SD). Differences between countries were tested using one-way analysis of variance (ANOVA).

Results

EVD case data

We used recent WHO datasets for Guinea, Liberia and Sierra of May 11, 2016, which contained 4391 data points for 56 affected of a total of 63 districts for the years 2013–2015. Three districts were excluded because they had fewer than three non-zero data points in one wave (Dinguiraye, Guinea; Togué, Guinea and Bonthe, Sierra Leone). The restriction of the dataset to the first 8 weeks since onset resulted in 477 data points for 53 districts (nine data points per district). For most districts, the first epidemic wave was large enough to be included in the analysis, except for seven districts where we used the second wave (Boké, Fria, Kindia and Siguiri in Guinea; Margibi and Nimba in Liberia and Koinadugu in Sierra Leone) and one district where we used the fourth wave (Kouroussa in Guinea). The different epidemic waves in each district are displayed in the supplement (S1 Fig). On visual inspection, the restriction of the time variable captured the initial growth phase appropriately (Fig 1).

Basic reproduction number R_0

Epidemic growth rates and thus R_0 differed at both district and country level. The spatial distribution of district-level estimates shows that districts with high transmissibility appear to cluster regionally irrespective of national borders (Fig 2). At district level, R_0 ranged from 0.36 (Dubréka) to 1.72 (Beyla) in Guinea, from 0.53 (Maryland) to 3.37 (Margibi) in Liberia and from 1.14 (Koinadugu) to 2.73 (Western Rural) in Sierra Leone (Table 1). Transmissibility was below the epidemic threshold of $R_0 = 1$ in 56% of all districts in Guinea, in 33% of all district in Liberia and in none of the districts in Sierra Leone (Fig 3). District-level R_0 values differed not more between than within the three countries (one-way ANOVA, $F(2, 50) = 8.38$, $p = 0.083$). The values of R_0 and 95% CIs for each district are provided in the supplement (S1 Table). We also estimated a national R_0 of 0.97 (95% CI 0.77–1.18) for Guinea, 1.26 (95% CI 0.98–1.55) for Liberia and 1.66 (95% CI 1.32–2.00) for Sierra Leone. The overall mean of the district-level R_0 values was 1.30 (SD 0.64) and the distribution was right-skewed (S2 Fig).

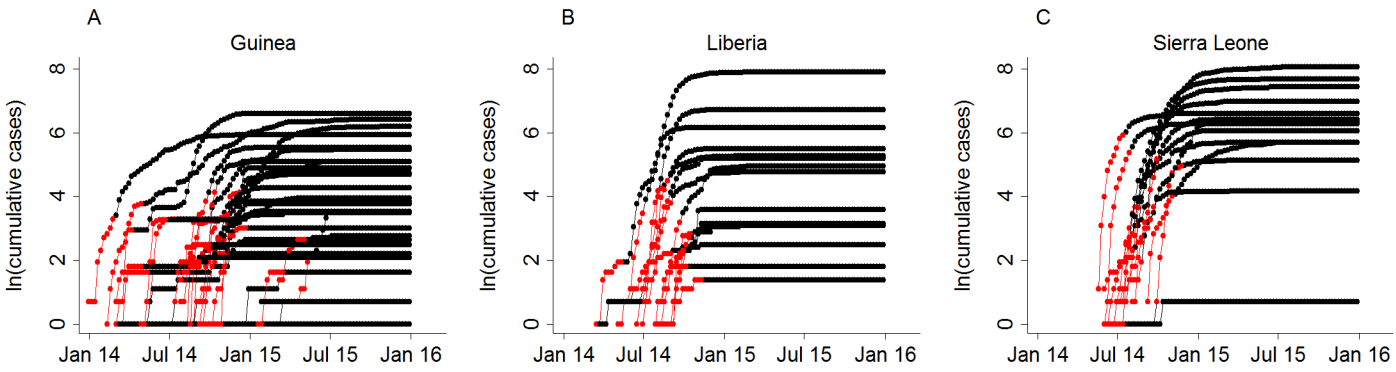

Fig 1. Natural log-transformed cumulative EVD case numbers by district and country. A. Guinea. **B**. Liberia. **C**. Sierra Leone. The red dots indicate the data points included in the mixed effects model.

doi:10.1371/journal.pntd.0004867.g001

The inclusion of shorter or longer time windows for the mixed effects model had little influence on the magnitude of the national estimates of r and R_0, but as expected the uncertainty in the estimate increased with decreasing length of the time window (S2 Table and S3 Fig). At district level, the median did not vary substantially in Guinea, but some districts showed more extreme estimates for time windows of five weeks or more than nine weeks. In Liberia and Sierra Leone the median decreased towards one for time windows of five weeks or more than ten weeks.

Associations between R_0 and socio-demographic factors

We found no association between R_0 and household density (β = 0.51, 95% CI -0.30–1.31, p = 0.215) or time taken to walk to the nearest water source (β = -0.18, 95% CI -0.43–0.07, p = 0.151) (Table 2 and Fig 4). There was weak statistical evidence for a positive association between log-transformed population density and R_0 (β = 0.12, 95% CI -0.02–0.26, p = 0.086). The DHS wealth index score showed the strongest statistical evidence for a positive association with R_0 (β = 0.37, 95% CI 0.07–0.67, p = 0.017). This result appears counterintuitive, but an increasing wealth score was also positively associated with population density suggesting that both variables act as a proxy for urbanisation and its effects on human contact patterns. These findings suggest that variables associated with large-scale crowding may act as population-level risk factors for EVD transmission.

Discussion

Our study provides further evidence that the EVD epidemic in West Africa was a spatially heterogeneous process at district level, and that socio-demographic factors might have contributed to the spread of EVD. Average district transmissibility was lower than the epidemic threshold of $R_0 = 1$ in Guinea but higher in Liberia and Sierra Leone. Geographically adjacent areas appeared to have a similar transmissibility regardless of country borders. The spatial distribution of transmission estimates suggested a cluster region in the coastal districts in the southwest of Sierra Leone and Liberia and in the east of Guinea and Liberia, and less intense transmission in the north of Guinea. Population density and a high DHS wealth index score at district-level were positively associated with R_0 suggesting that factors related to urbanisation and large-scale crowding might have contributed to the rapid spread of EVD in certain areas.

The strength of our statistical approach is that it allowed for a more realistic scenario of growth rates than models that treat each outbreak individually. The use of a generalised linear

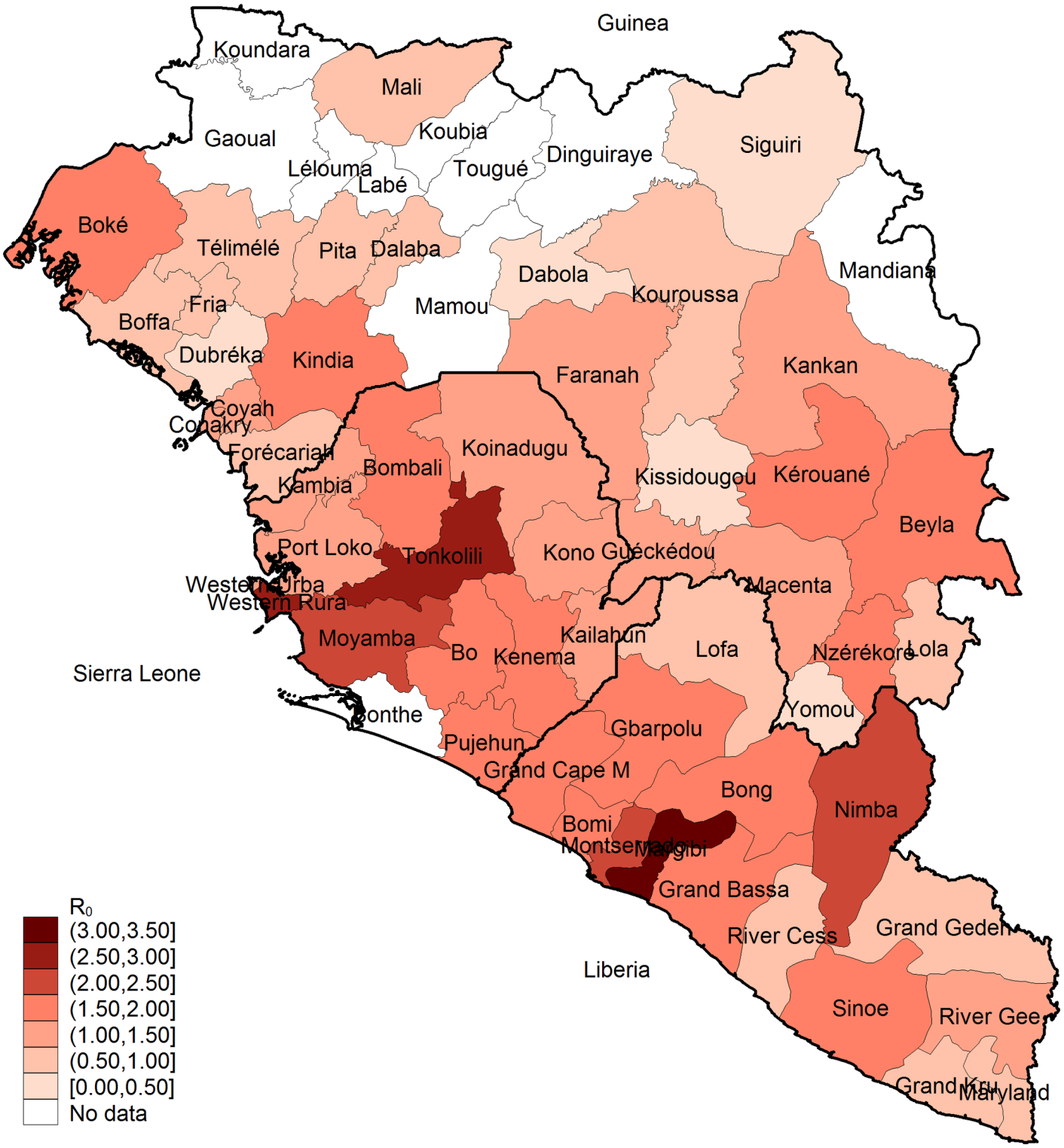

Fig 2. Geographical distribution of district-level R_0. The shapefiles were retrieved from the Database of global administrative areas GADM ([45]).

doi:10.1371/journal.pntd.0004867.g002

Table 1. District-level and national estimates of R_0.

	Guinea (N = 25)		Liberia (N = 15)		Sierra Leone (N = 13)	
Districts†	**Median**	**(range)**	**Median**	**(range)**	**Median**	**(range)**
R_0	0.91	(0.36–1.72)	1.68	(0.53–3.37)	1.50	(1.14–2.73)
National*	**Estimate**	**(95% CI)**	**Estimate**	**(95% CI)**	**Estimate**	**(95% CI)**
R_0	0.97	(0.77–1.18)	1.26	(0.98–1.55)	1.66	(1.32–2.00)

† The district-level growth rate estimates correspond to the sum of the country-specific fixed effect and the district-specific random effect

* The national growth rate estimates correspond to the fixed effects of the GLMM

doi:10.1371/journal.pntd.0004867.t001

mixed model provides a simple but elegant solution for a geographically complex and hierarchical epidemic structure and yielded plausible values of transmissibility. As shown in our sensitivity analysis, the length of the time window for the exponential phase did not affect the estimates of R_0 substantially. The selection of the time interval to be included should be based on considerations about the generation time, the number of data points available and knowledge about the time point of control interventions. Our approach has three main limitations, two related to data quality and one to the choice of methods. First, the reliability of the data collected at the beginning of the epidemic is uncertain [46]. Reporting delays and underreporting could have led to the underestimation of the incidence. Delays in updating the patient database do not affect our analysis because we included only the initial data points for each district. Underreporting of cases was estimated to be 17 to 70% [47]. If underreporting remains constant throughout the observed time period, R_0 is unaffected, because the exponential growth rate does not change. The assumption of proportional underreporting seems reasonable for the initial phase of two months even if it does not hold throughout the epidemic. Due to the lack of data on dynamic underreporting, we did not consider this aspect in our analysis. Second, our assumption about the absence of control measures might not be true. Local outbreaks occurred at different time points of the epidemic. Immediate implementation of control measures or increased public awareness in later outbreaks could have biased estimates of the initial epidemic growth downwards for districts affected towards the end of 2014. We think that the effect of these interventions was negligible during the first two months of an outbreak because, even by early October 2014, only a few districts had managed to implement fully functioning control measures such as safe burials and contact tracing [48] and most ETUs opened more than eight weeks after onset of the epidemic in a district. Third, our approach does not take into account the spatial dependence of cases and in theory the geographical location of cases is exchangeable within a country in our model. Spatial autocorrelation models require a spatial weights matrix, which is derived using geographical information, and are conceptually more complex than normal regression models. We aimed to establish a simple model, which can be used when limited geographical information is available but which still includes the hierarchical aspect of spatially heterogeneous epidemics. This analysis also has a limited spatial resolution due to the availability of data and cannot capture heterogeneities at a finer scale. The assumption of homogeneous population mixing within a district is clearly unrealistic. However, the trade-off between higher spatial resolution and lower case numbers is not straightforward. Chiefdom-level data may provide a better picture of local dynamics but are not publicly available.

Our district-level estimates of R_0 are generally consistent with results of studies from Kenema (Sierra Leone), Montserrado (Liberia) and Conakry (Guinea) (S3 Table). Four studies using different phylodynamic models based on patient samples taken in a hospital in Kenema have given estimates of R_0 between 1.26 and 2.40 [23–25,47]. Our estimate of 1.65 (95% CI

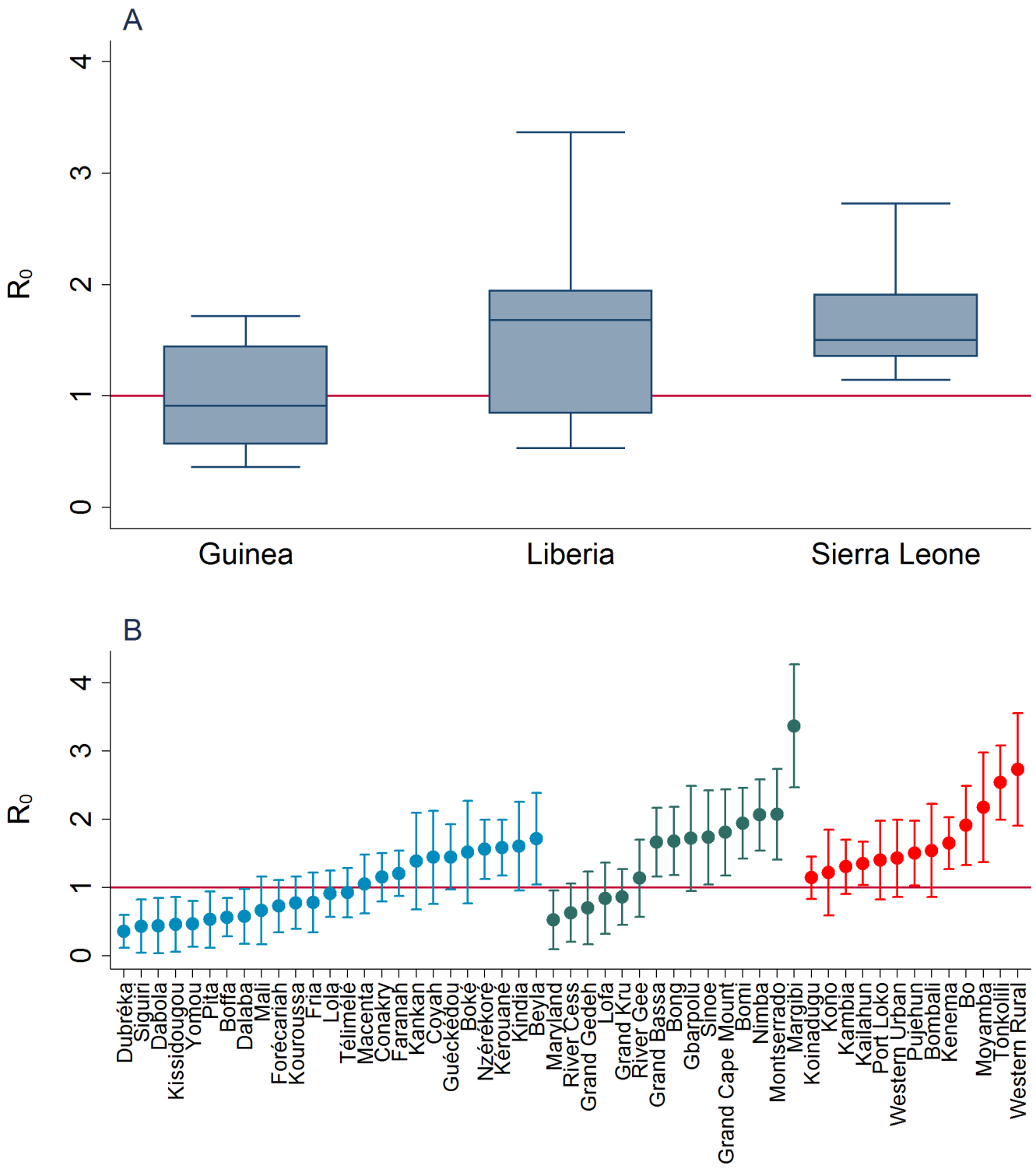

Fig 3. District-level estimates of R_0. The red horizontal line indicates the epidemic threshold of $R_0 = 1$. **A**. The box shows the interquartile range, the horizontal line is the country median. The ends of the whiskers are the lower and upper range values. **B**. Individual estimates and 95% confidence intervals of R_0 for districts in Guinea (blue), Liberia (green) and Sierra Leone (red).

doi:10.1371/journal.pntd.0004867.g003

Table 2. Association between socio-demographic exposure variables and R_0 in a linear univariable model.

Variable	Unit	β	(95% CI)	p-value
Population density	per $\log_e$	0.12	(-0.02–0.26)	0.086
Household density	per person	0.51	(-0.30–1.31)	0.215
DHS wealth score	per 10^5	0.37	(0.07–0.67)	0.017
Time to water source	per 10 minutes	-0.18	(-0.43–0.07)	0.151

DHS Demographic and Health Survey, p-value from Wald test

doi:10.1371/journal.pntd.0004867.t002

1.27–2.03) is within this range. For Montserrado, we estimated an R_0 of 2.07 (95% CI 1.41–2.74), which lies between the published estimates of 1.73 and 2.49 [21,22]. For Conakry, our model yielded a lower value (1.15, 95% CI 0.80–1.51) than that calculated from reconstructed transmission chains (1.7) for the time period of March 2014 [20] but was close to the overall value of 0.95 estimated from a re-analysis of the same data [49]. The finding that many districts

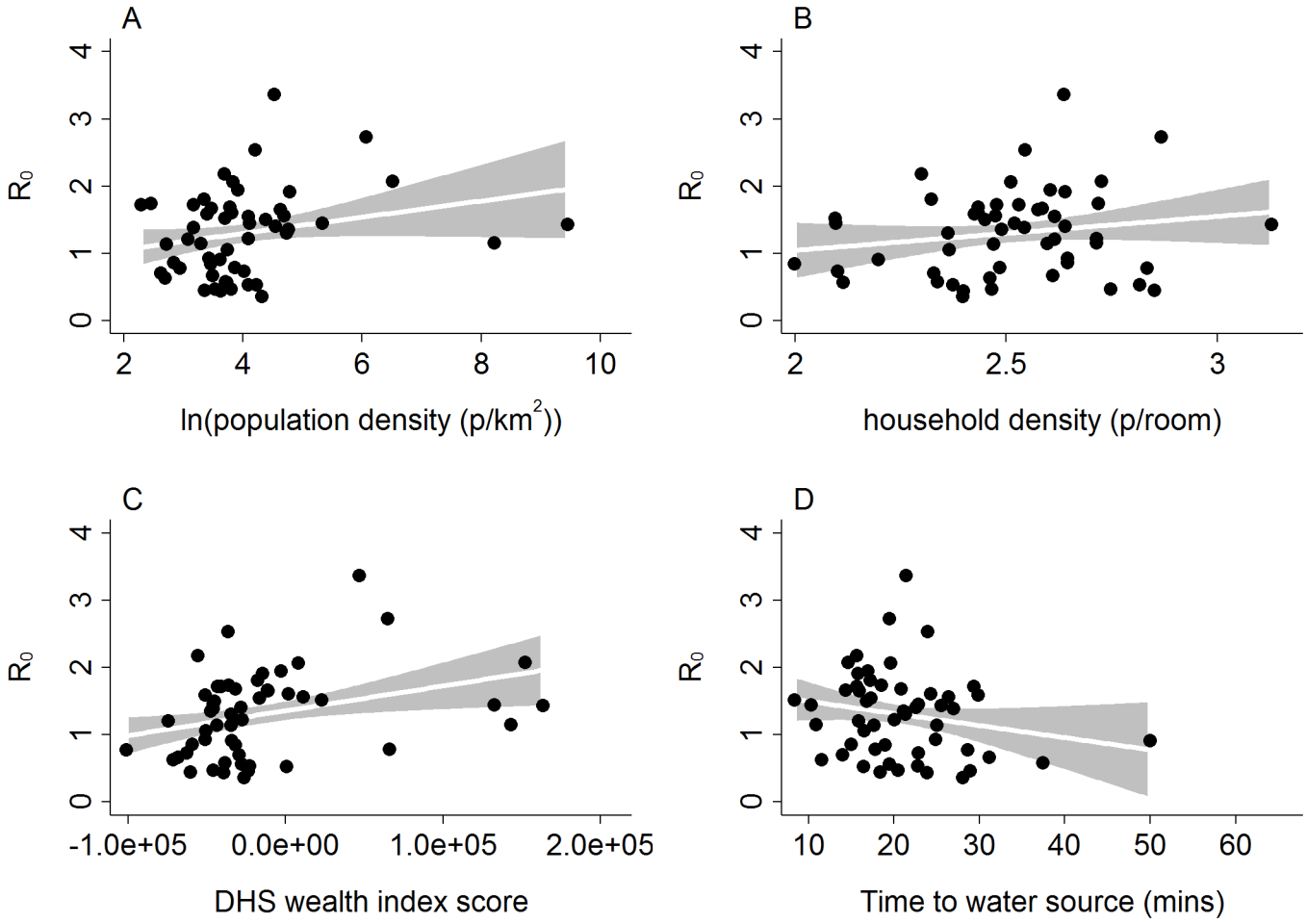

Fig 4. Scatterplots of R_0 and socio-demographic exposure variables. A. Association between R_0 and log-transformed population density (p/km^2). **B**. Association between R_0 and household density (p/room). **C**. Association between R_0 and DHS wealth index score. **D**. Association between R_0 and time to water source (mins). The white line denotes the linear fit, the grey shaded area denotes the 95% CI. DHS = Demographic and Health Survey, p = persons.

doi:10.1371/journal.pntd.0004867.g004

experienced an outbreak that was characterised by an $R_0<1$ supports the hypothesis of heterogeneous contact networks [4,11]. Clustered transmission and superspreading has been confirmed by several studies [20,25,47] and can lead to outbreaks even if R_0 is below unity [49]. Assuming that our approach is an adequate model of this hierarchical epidemic, we may conclude that subnational transmissibility might not have been as high as previously thought. At national level, our approach was not comparable to results of other methods. Several mechanistic [6,10,14,16–19,50–52] and phenomenological [15,34,53] models have provided estimates of R_0 ranging from 1.51 (95% CI 1.50–1.52) to 2.46 (95% CI 1.44–2.01) for Guinea, from 1.54 to 2.5 (95% CI 2.4–2.7) for Liberia and from 1.26 to 8.33 for Sierra Leone. Our estimates of 0.97 (95% CI 0.77–1.18) for Guinea, of 1.26 (95% CI 0.98–1.55) for Liberia and of 1.66 (95% CI 1.43–2.00) for Sierra Leone are consistently lower than all published values. These differences might reflect the fact that our national estimates of R_0 result from an averaged growth rate of multiple local outbreaks occurring at different time points, whilst other studies used the national epidemic curves to fit their models.

The ecological analysis showed associations between R_0 and population-level factors linked to urbanisation and crowding. Analysis at the population-level is appropriate when the mechanism of action involves interactions between individuals, for example the potential for spread of an infection in densely populated areas [54]. Nevertheless, the risk of an individual of becoming infected cannot be predicted from the population density of their household or community. Whether the positive association between crowding and EVD transmissibility is also observed at a lower level of administrative unit cannot be inferred from this analysis. Fallah et al used individual data on EVD cases and their contacts and determined SES-stratified measures of transmission [12]. They found that cases from middle and low SES communities caused significantly more secondary cases than infected individuals from high SES communities. These findings appear contradictory to our study, but can be explained by the different levels of data and different definitions of the SES. While our SES variable considers housing properties and household assets, Fallah et al also included high population density in the definition of low SES. Considering these differences, both studies are consistent with the hypothesis that large-scale crowding contributed to EVD transmission. Other factors that could influence the transmissibility could not be considered in this analysis due to lack of data. Behavioural factors such as community resistance or superspreading events like unsafe burials [55] might have had a strong impact on EVD spread. Some of our estimates may be large by chance due to the stochastic nature of the outbreaks. Biological factors such as differences in host or viral genetics are probably less important in this population and this time frame. Our findings are compatible with theory about drivers of infectious diseases, but the contribution of different factors cannot be answered completely with this study design and these data.

This study has shown that the mixed effects model is a suitable strategy to quantify local epidemic growth during a large-scale multifocal epidemic and confirms the notion of geographical heterogeneity in the transmission of EVD in West Africa. Social and demographic variables measured at the population level and related to urbanisation, such as high population density and high SES, were positively associated with R_0 suggesting that the consequences of fast urban growth in West Africa may have contributed to the increased spread of EVD.

Supporting Information

S1 Text. Calculation of summary estimates from DHS datasets.
(PDF)

S1 Fig. Epidemic curves with distinct epidemic waves for each district. The colors denote the order of the waves: red = first wave, blue = second wave, green = third wave,

magenta = fourth wave, black = interwave periods with no cases
(TIF)

S2 Fig. Overall distribution of district-level R_0. The histogram was calculated with a bin width of 0.1.
(TIF)

S3 Fig. Influence of different time windows included in the mixed effects model on magnitude of R_0. A-C. National estimates (line) and 95% confidence intervals (shaded area). **D-F.** District-level medians (line) and range (shaded area).
(TIF)

S1 Table. Individual estimates and confidence intervals of R_0 for each district.
(XLS)

S2 Table. Sensitivity analysis with varying time windows for national and district-level R_0.
(PDF)

S3 Table. Comparison of other published estimates (95% confidence or credible intervals) of R_0
(PDF)

Acknowledgments

We thank the reviewers for their constructive comments and suggestions. We also thank all persons who were involved in the outbreak response for collecting, cleaning and publishing EVD case data under difficult circumstances.

Author Contributions

Conceived and designed the experiments: FK SG NL CHH CLA. Analyzed the data: FK. Wrote the paper: FK SG NL CHH CLA.

References

1. Towers S, Patterson-lomba O, Castillo-chavez C. Temporal Variations in the Effective Reproduction Number of the 2014 West Africa Ebola Outbreak. PLOS Curr Outbreaks. 2014;Sep 18(1).
2. Plachouras D, Sudre B, Testa M, Robesyn E, Coulombier D. Letter to the editor: Early transmission dynamics of Ebola virus disease (EVD), West Africa, March to August 2014 –Eurosurveillance 17 September 2014. Eurosurveillance. 2014 Sep 18; 19(37):20907.
3. Nishiura H, Chowell G. Authors ' reply : Feedback from modelling to surveillance of Ebola virus disease. Eurosurveillance. 2014;Sep 18(19(37)).
4. Chowell G, Viboud C, Hyman JM, Simonsen L. The Western Africa ebola virus disease epidemic exhibits both global exponential and local polynomial growth rates. PLoS Curr. 2015; 7(1).
5. Santermans E, Robesyn E, Ganyani T, Sudre B, Faes C, Quinten C, et al. Spatiotemporal Evolution of Ebola Virus Disease at Sub-National Level during the 2014 West Africa Epidemic: Model Scrutiny and Data Meagreness. PLoS One. 2016; 11(1):e0147172. doi: 10.1371/journal.pone.0147172 PMID: 26771513
6. Merler S, Ajelli M, Fumanelli L, Gomes MFC, Piontti AP y, Rossi L, et al. Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. Lancet Infect Dis. 2015 Feb; 15(2):204–11. doi: 10.1016/S1473-3099(14)71074-6 PMID: 25575618
7. Fang L-Q, Yang Y, Jiang J-F, Yao H-W, Kargbo D, Li X-L, et al. Transmission dynamics of Ebola virus disease and intervention effectiveness in Sierra Leone. Proc Natl Acad Sci U S A. 2016 Apr 19; 113 (16):4488–93. doi: 10.1073/pnas.1518587113 PMID: 27035948

8. Kucharski AJ, Camacho A, Flasche S, Glover RE, Edmunds WJ, Funk S. Measuring the impact of Ebola control measures in Sierra Leone. Proc Natl Acad Sci U S A. 2015 Nov 17; 112(46):14366–71. doi: 10.1073/pnas.1508814112 PMID: 26460023

9. Kiskowski MA. A three-scale network model for the early growth dynamics of 2014 west Africa ebola epidemic. PLoS Curr. 2014; 6(1).

10. White RA, MacDonald E, de Blasio BF, Nygård K, Vold L, Røttingen J-A. Projected Treatment Capacity Needs in Sierra Leone. PLoS Curr. 2014;1–16.

11. Chowell G, Nishiura H. Characterizing the transmission dynamics and control of ebola virus disease. PLoS Biol. 2015 Jan; 13(1):e1002057. doi: 10.1371/journal.pbio.1002057 PMID: 25607595

12. Fallah MP, Skrip LA, Gertler S, Yamin D, Galvani AP. Quantifying Poverty as a Driver of Ebola Transmission. PLoS Negl Trop Dis. 2015 Dec; 9(12):e0004260. doi: 10.1371/journal.pntd.0004260 PMID: 26720278

13. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci. 2007; 274:599–604. PMID: 17476782

14. Althaus CL. Estimating the Reproduction Number of Ebola Virus (EBOV) During the 2014 Outbreak in West Africa. PLoS Curr. 2014; 6(1):1–9.

15. Fisman D, Khoo E, Tuite A. Early epidemic dynamics of the west african 2014 ebola outbreak: estimates derived with a simple two-parameter model. PLoS Curr. 2014; 6(1).

16. Rivers C, Lofgren E, Marathe M, Eubank S, Lewis B. Modeling the Impact of Interventions on an Epidemic of Ebola in Sierra Leone and Liberia. PLoS Curr. 2014;Oct 16(1).

17. Webb G, Browne C, Huo X, Seydi O, Seydi M, Magal P. A Model of the 2014 Ebola Epidemic in West Africa with Contact Tracing. PLoS Curr. 2015;1–10.

18. Khan A, Naveed M, Dur-E-Ahmad M, Imran M. Estimating the basic reproductive ratio for the Ebola outbreak in Liberia and Sierra Leone. Infect Dis poverty. London; 2015; 4(13):13.

19. Siettos C, Anastassopoulou C, Russo L, Grigoras C, Mylonakis E. Modeling the 2014 Ebola Virus Epidemic—Agent-Based Simulations, Temporal Analysis and Future Predictions for Liberia and Sierra Leone. PLoS Curr. 2015;(December 2013):1–16.

20. Faye O, Boelle PY, Heleze E, Faye O, Loucoubar C, Magassouba N, et al. Chains of transmission and control of Ebola virus disease in Conakry, Guinea, in 2014: an observational study. Lancet Infect Dis. 2015;Jan 22(14):1–7.

21. Yamin D, Gertler S, Ndeffo-Mbah ML, Skrip LA, Fallah M, Nyenswah TG, et al. Effect of Ebola progression on transmission and control in Liberia. Ann Intern Med. 2015 Jan 6; 162(1):11–7. doi: 10.7326/M14-2255 PMID: 25347321

22. Lewnard JA, Ndeffo Mbah ML, Alfaro-Murillo JA, Altice FL, Bawo L, Nyenswah TG, et al. Dynamics and control of Ebola virus transmission in Montserrado, Liberia: a mathematical modelling analysis. Lancet Infect Dis. Oxford; 2014; 14(12):1189–95.

23. Stadler T, Kühnert D, Rasmussen DA, du Plessis L. Insights into the early epidemic spread of ebola in sierra leone provided by viral sequence data. PLoS Curr. 2014; 6.

24. Alizon S, Lion S, Murall CL, Abbate JL. Quantifying the epidemic spread of Ebola virus (EBOV) in Sierra Leone using phylodynamics. Virulence. 2014; 5(8):825–7. doi: 10.4161/21505594.2014.976514 PMID: 25495064

25. Volz E, Pond S. Phylodynamic analysis of ebola virus in the 2014 sierra leone epidemic. PLoS Curr. 2014; 6.

26. Van Kerkhove MD, Bento AI, Mills HL, Ferguson NM, Donnelly CA. A review of epidemiological parameters from Ebola outbreaks to inform early public health decision-making. Sci data. 2015; 2:150019. doi: 10.1038/sdata.2015.19 PMID: 26029377

27. Ma J, Dushoff J, Bolker BM, Earn DJD. Estimating initial epidemic growth rates. Bull Math Biol. 2014 Jan; 76(1):245–60. doi: 10.1007/s11538-013-9918-2 PMID: 24272389

28. Vynnycky E, White R. An Introduction to Infectious Disease Modelling. An Introduction to Infectious Disease Modelling. Oxford University Press; 2010. 400 p.

29. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003 Jun 20; 300(5627):1966–70. PMID: 12766207

30. de Silva UC, Warachit J, Waicharoen S, Chittaganpitch M. A preliminary analysis of the epidemiology of influenza A(H1N1)v virus infection in Thailand from early outbreak data, June-July 2009. Euro Surveill Bull Eur sur les Mal Transm = Eur Commun Dis Bull. 2009 Aug 6; 14(31).

31. King AA, de Celles MD, Magpantay FMG, Rohani P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. Proc R Soc London Ser B, Biol Sci. London; 2015; 282(1806):20150347.

32. World Health Organisation WHO. Ebola data and statistics [Internet]. Available from: http://apps.who.int/gho/data/node.ebola-sitrep.ebola-country?lang=en

33. Camacho A, Kucharski A, Aki-Sawyerr Y, White MA, Flasche S, Baguelin M, et al. Temporal Changes in Ebola Transmission in Sierra Leone and Implications for Control Requirements: a Real-time Modelling Study. PLoS Curr. 2015; 7(Grant 13165):1–19.

34. WHO Ebola Response Team. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. N Engl J Med. United States; 2014 Sep 22; 371(16):1481–95.

35. Exchange HD. West Africa: Ebola Outbreak [Internet]. [cited 2016 Mar 1]. Available from: https://data.humdata.org/dataset/ebola-treatment-centers

36. World Health Organization WHO. Overview of Ebola Treatment Centres (ETCs) [Internet]. [cited 2015 Jul 31]. Available from: https://extranet.who.int/ebolafmt/sites/default/files/documents/ETC_Overview_05_April_2015.pdf

37. Kirkwood BB, Sterne J. Essential medical statistics. Blackwell Publishing; 2003. 1–512 p.

38. Institut National de la Statistique G. Résultats préliminaires du RGPH3 [Internet]. 2014 [cited 2015 Jan 31]. Available from: http://www.stat-guinee.org/index.php/result-prelim-rgph3

39. Liberian Institute of Statistics and Geo-Information Services (LISGIS). 2008 National Population and Housing Census. Analytical report on population projections [Internet]. 2011. Available from: http://www.lisgis.net/page_info.php?7d5f44532cbfc489b8db9e12e44eb820=MzQy

40. Koroma DS, Turay AB, Moigua MB. 2004 Population and Housing Census. Analytical report on population projection for Sierra Leone [Internet]. 2006 [cited 2015 Jan 1]. Available from: http://www.statistics.sl/2004_population_and_housing_census.htm

41. Institut National de la Statistique (INS) et ICF International. L'Enquête Démographique et de Santé et à Indicateurs Multiples en Guinée de 2012 [Dataset]. 2013.

42. Liberia Institute of Statistics and GeoInformation Services (LISGIS), Ministry of Health and Social Welfare, National AIDS Control Program and ICF International. Liberia Demographic and Health Survey 2013 [Dataset]. 2014.

43. Statistics Sierra Leone (SSL) and ICF International. Sierra Leone Demographic and Health Survey 2013 [Dataset]. 2014.

44. Rutstein S, Staveteig S. Making the Demographic and Health Surveys Wealth Index Comparable. DHS Methodological Reports No. 9. Rockville, Maryland, USA: ICF International.; 2014.

45. GADM database of global administrative areas, Version 2.5 [Internet]. 2012. Available from: http://www.gadm.org

46. World Health Organization WHO. Factors that contributed to undetected spread of the Ebola virus and impeded rapid containment [Internet]. 2015 [cited 2015 Dec 29]. Available from: http://www.who.int/csr/disease/ebola/one-year-report/factors/en/

47. Scarpino S V, Iamarino A, Wells C, Yamin D, Ndeffo-Mbah M, Wenzel NS, et al. Epidemiological and viral genomic sequence analysis of the 2014 ebola outbreak reveals clustered transmission. Clin Infect Dis. 2015; 60(7):1079–82. doi: 10.1093/cid/ciu1131 PMID: 25516185

48. World Health Organisation WHO. Ebola Response Roadmap Situation Report [Internet]. 2014. Available from: http://www.who.int/csr/disease/ebola/situation-reports/en/

49. Althaus CL. Ebola superspreading. Lancet Infect Dis. 2015 May; 15(5):507–8. doi: 10.1016/S1473-3099(15)70135-0 PMID: 25932579

50. Pandey A, Atkins KE, Medlock J, Wenzel N, Townsend JP, Childs JE, et al. Strategies for containing Ebola in West Africa. Science. 2014 Nov 21; 346(6212):991–5. doi: 10.1126/science.1260612 PMID: 25414312

51. Valdez LD, Aragão Rêgo HH, Stanley HE, Braunstein LA. Predicting the extinction of Ebola spreading in Liberia due to mitigation strategies. Sci Rep. 2015; 5:12172. doi: 10.1038/srep12172 PMID: 26190582

52. Xia Z-Q, Wang S-F, Li S-L, Huang L-Y, Zhang W-Y, Sun G-Q, et al. Modeling the transmission dynamics of Ebola virus disease in Liberia. Sci Rep. 2015; 5:13857. doi: 10.1038/srep13857 PMID: 26347015

53. Chowell G, Nishiura H. Transmission dynamics and control of Ebola virus disease (EVD): a review. BMC Med. 2014 Jan 21; 12(1):196.

54. Susser M. The logic in ecological: I. The logic of analysis. Am J Public Health. 1994 May; 84(5):825–9. PMID: 8179056

55. Victory KR, Coronado F, Ifono SO, Soropogui T, Dahl BA, Centers for Disease Control and Prevention (CDC). Ebola transmission linked to a single traditional funeral ceremony—Kissidougou, Guinea, December, 2014-January 2015. MMWR Morb Mortal Wkly Rep. Atlanta; 2015; 64(14):386–8.